



(translation of the front page of the priority document of  
Japanese Patent Application No. 2001-118348)

PATENT OFFICE  
JAPANESE GOVERNMENT

This is to certify that the annexed is a true copy of the  
following application as filed with this Office.

Date of Application: April 17, 2001

Application Number : Patent Application 2001-118348

Applicant(s) : Canon Kabushiki Kaisha

May 18, 2001

Commissioner,  
Patent Office

Kouzo OIKAWA

Certification Number 2001-3041092

CFM 2209 US  
09/842,155



日 本 国 特 許 庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日  
Date of Application:

2001年 4月17日

出 願 番 号  
Application Number:

特願2001-118348

出 願 人  
Applicant(s):

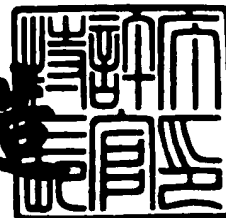
キヤノン株式会社

CERTIFIED COPY OF  
PRIORITY DOCUMENT

2001年 5月18日

特 許 庁 長 官  
Commissioner,  
Japan Patent Office

及 川 耕 造



出証番号 出証特2001-3041092

【書類名】 特許願

【整理番号】 4436011

【提出日】 平成13年 4月17日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 7/00

【発明の名称】 符号化装置及び符号化方法

【請求項の数】 23

【発明者】

    【住所又は居所】 東京都大田区下丸子3丁目30番2号 キヤノン株式会社  
                        社内

    【氏名】 岸 裕樹

【特許出願人】

    【識別番号】 000001007

    【氏名又は名称】 キヤノン株式会社

【代理人】

    【識別番号】 100076428

    【弁理士】

    【氏名又は名称】 大塚 康德

    【電話番号】 03-5276-3241

【選任した代理人】

    【識別番号】 100112508

    【弁理士】

    【氏名又は名称】 高柳 司郎

    【電話番号】 03-5276-3241

【選任した代理人】

    【識別番号】 100115071

    【弁理士】

    【氏名又は名称】 大塚 康弘

    【電話番号】 03-5276-3241

【選任した代理人】

【識別番号】 100116894

【弁理士】

【氏名又は名称】 木村 秀二

【電話番号】 03-5276-3241

【先の出願に基づく優先権主張】

【出願番号】 特願2000-128515

【出願日】 平成12年 4月27日

【手数料の表示】

【予納台帳番号】 003458

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 0102485

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 符号化装置及び符号化方法

【特許請求の範囲】

【請求項 1】 画像データと音声データを含むフレームデータに対して符号化を行う符号化装置であって、

前記フレームデータに含まれる前記画像データと前記音声データを分離する分離手段と、

分離された前記画像データに対し、その低周波成分から高周波成分へ順次、符号化を行って画像符号化データを生成する画像データ符号化手段と、

分離された前記音声データに対し、その低周波成分から高周波成分へ順次、符号化を行って音声符号化データを生成する音声データ符号化手段と、

前記画像符号化データと前記音声符号化データを用いてヘッダ情報を生成し、該ヘッダ情報及び前記画像符号化データ及び前記音声符号化データによってフレーム符号化データを生成するフレーム符号化データ生成手段と、  
を備えることを特徴とする符号化装置。

【請求項 2】 前記ヘッダ情報は少なくとも、前記画像データのサイズ、前記画像データの画像タイプ情報、前記画像符号化データの長さ、前記音声符号化データの長さ、当該符号化装置の識別情報、送信日時、前記画像符号化データの先頭アドレス、前記音声符号化データの先頭アドレス、のいずれかを含むことを特徴とする請求項 1 記載の符号化装置。

【請求項 3】 前記画像データ符号化手段は、前記画像データに対して離散ウェーブレット変換を施すことによって各サブバンドに対する変換係数列を生成し、同レベルのサブバンドをまとめて、低レベルのサブバンドから高レベルのサブバンドへ順次、前記変換係数列を符号化することによって、前記画像符号化データを生成することを特徴とする請求項 1 記載の符号化装置。

【請求項 4】 前記音声データ符号化手段は、前記音声データに対して離散ウェーブレット変換を施すことによって各サブバンドに対する変換係数列を生成し、同レベルのサブバンドをまとめて、低レベルのサブバンドから高レベルのサブバンドへ順次、前記変換係数列を符号化することによって、前記音声符号化デ

ータを生成することを特徴とする請求項 1 記載の符号化装置。

【請求項 5】 前記フレーム符号化データ生成手段は、前記ヘッダ情報、前記画像符号化データ、前記音声符号化データをこの順に並べることによって、前記フレーム符号化データを生成することを特徴とする請求項 1 乃至 4 のいずれかに記載の符号化装置。

【請求項 6】 前記フレーム符号化データ生成手段は、前記画像符号化データと前記音声符号化データにおける同レベルのサブバンドをグループ化し、該グループを前記ヘッダ情報に続けて低レベルから高レベルの順に並べることによって、前記フレーム符号化データを生成することを特徴とする請求項 1 乃至 4 のいずれかに記載の符号化装置。

【請求項 7】 前記フレーム符号化データ生成手段は、前記画像符号化データの一部と前記音声符号化データの一部から構成される準フレーム符号化データを用いて、前記フレーム符号化データを生成することを特徴とする請求項 1 乃至 4 のいずれかに記載の符号化装置。

【請求項 8】 画像データと音声データを含むフレームデータに対して符号化を行う符号化装置であって、

前記フレームデータに含まれる前記画像データと前記音声データを分離する分離手段と、

前記画像データを複数種類の画像データに階層化して符号化し、複数階層の夫々に相当する符号化画像データを生成する画像データ符号化手段と、

前記音声データを複数種類の音声データに階層化して符号化し、複数階層の夫々に相当する符号化音声データを生成する音声データ符号化手段と、

前記符号化画像データと前記符号化音声データを用いてフレーム符号化データを生成するフレーム符号化データ生成手段と、を備え、

前記フレーム符号化データ生成手段は、所定基準に基づいて決定される同一レベルに属する画像符号化データと音声符号化データをグループ化することによってレベルの異なる複数のグループを作成し、該複数のグループを重要度の高いレベル順に並べることによって前記フレーム符号化データを生成することを特徴とする符号化装置。

【請求項 9】 前記画像データ符号化手段が階層化した複数種類の画像データは、前記画像データを離散ウェーブレット変換することによって得られる複数の周波数成分に相当することを特徴とする請求項 8 記載の符号化装置。

【請求項 10】 前記音声データ符号化手段が階層化した複数種類の音声データは、人間の声に相当するスピーチデータと、それ以外の非スピーチデータに相当することを特徴とする請求項 8 記載の符号化装置。

【請求項 11】 前記フレーム符号化データ生成手段は、前記スピーチデータの符号化データを重要レベルの音声符号化データとして第 1 の符号化画像データとグループ化し、前記非スピーチデータの符号化データを非重要レベルの音声符号化データとして第 2 の符号化画像データとグループ化することを特徴とする請求項 10 記載の符号化装置。

【請求項 12】 前記画像データ符号化手段が階層化した複数種類の画像データは、前記画像データを離散ウェーブレット変換することによって得られる第 1 の周波数成分と、該第 1 の周波数成分よりも高周波である第 2 の周波数成分を含み、

前記第 1 及び第 2 の符号化画像データは夫々、該第 1 及び第 2 の周波数成分に相当する符号化画像データであることを特徴とする請求項 11 記載の符号化装置。

【請求項 13】 前記音声データ符号化手段が階層化した複数種類の音声データは、人間の声に相当するスピーチデータと、それ以外の非スピーチデータを階層化した 2 つ以上の非スピーチデータに相当することを特徴とする請求項 8 記載の符号化装置。

【請求項 14】 前記フレーム符号化データ生成手段は、  
前記スピーチデータの符号化データを最重要レベルの音声符号化データとして第 1 の符号化画像データとグループ化し、

前記非スピーチデータを階層化して得られた第 1 非スピーチデータの符号化データを、前記最重要レベルに次いで重要なレベルの音声符号化データとして第 2 の符号化画像データとグループ化し、

前記非スピーチデータを階層化して得られた前記第 1 非スピーチデータ以外の

第 2 非スピーチデータの符号化データを、第 3 の符号化画像データとグループ化する

ことを特徴とする請求項 1 2 記載の符号化装置。

【請求項 1 5】 前記画像データ符号化手段が階層化した複数種類の画像データは、前記画像データを離散ウェーブレット変換することによって得られる第 1 の周波数成分と、該第 1 の周波数成分よりも高周波である第 2 の周波数成分と、該第 2 の周波数成分よりも高周波である第 3 の周波数成分を含み、

前記第 1 乃至第 3 の符号化画像データは夫々、該第 1 乃至第 3 の周波数成分に相当する符号化画像データであることを特徴とする請求項 1 4 記載の符号化装置。

【請求項 1 6】 前記フレーム符号化データ生成手段は、前記符号化画像データと前記符号化音声データをグループ化する際に、複数種類のグループ化方法を選択的に使用可能とすることを特徴とする請求項 8 記載の符号化装置。

【請求項 1 7】 前記複数種類のグループ化方法は、画質を優先したグループ化方法と、音質を優先したグループ化方法を含むことを特徴とする請求項 1 6 記載の符号化装置。

【請求項 1 8】 更に、前記フレーム符号化データを伝送する伝送手段と、該伝送したフレーム符号化データの復号状況を検知する検知手段と、該検知した復号状況に応じて、前記グループ化方法を切り替える制御手段と、を有することを特徴とする請求項 1 6 記載の符号化装置。

【請求項 1 9】 画像データと音声データを含むフレームデータに対して符号化を行う符号化方法であって、

前記フレームデータに含まれる前記画像データと前記音声データを分離する分離工程と、

分離された前記画像データに対し、その低周波成分から高周波成分へ順次、符号化を行って画像符号化データを生成する画像データ符号化工程と、

分離された前記音声データに対し、その低周波成分から高周波成分へ順次、符号化を行って音声符号化データを生成する音声データ符号化工程と、

前記画像符号化データと前記音声符号化データを用いてヘッダ情報を生成し、



該ヘッダ情報及び前記画像符号化データ及び前記音声符号化データによってフレーム符号化データを生成するフレーム符号化データ生成工程と、  
を備えることを特徴とする符号化方法。

【請求項20】 画像データと音声データを含むフレームデータに対して符号化を行う符号化方法であって、

前記フレームデータに含まれる前記画像データと前記音声データを分離する分離工程と、

前記画像データを複数種類の画像データに階層化して符号化し、複数階層の夫々に相当する符号化画像データを生成する画像データ符号化工程と、

前記音声データを複数種類の音声データに階層化して符号化し、複数階層の夫々に相当する符号化音声データを生成する音声データ符号化工程と、

前記符号化画像データと前記符号化音声データを用いてフレーム符号化データを生成するフレーム符号化データ生成工程と、を備え、

前記フレーム符号化データ生成工程においては、所定基準に基づいて決定される同一レベルに属する画像符号化データと音声符号化データをグループ化することによってレベルの異なる複数のグループを作成し、該複数のグループを重要度の高いレベル順に並べることによって前記フレーム符号化データを生成することを特徴とする符号化方法。

【請求項21】 コンピュータによって実行されることにより、画像データと音声データを含むフレームデータに対して符号化を行う符号化装置として該コンピュータを機能させるプログラムであって、

前記フレームデータに含まれる前記画像データと前記音声データを分離する分離工程のコードと、

分離された前記画像データに対し、その低周波成分から高周波成分へ順次、符号化を行って画像符号化データを生成する画像データ符号化工程のコードと、

分離された前記音声データに対し、その低周波成分から高周波成分へ順次、符号化を行って音声符号化データを生成する音声データ符号化工程のコードと、

前記画像符号化データと前記音声符号化データを用いてヘッダ情報を生成し、該ヘッダ情報及び前記画像符号化データ及び前記音声符号化データによってフレ

ーム符号化データを生成するフレーム符号化データ生成工程のコードと、  
を備えることを特徴とするプログラム。

【請求項 2 2】 コンピュータによって実行されることにより、画像データ  
と音声データを含むフレームデータに対して符号化を行う符号化装置として該コ  
ンピュータを機能させるプログラムであって、

前記フレームデータに含まれる前記画像データと前記音声データを分離する分  
離工程のコードと、

前記画像データを複数種類の画像データに階層化して符号化し、複数階層の夫  
々に相当する符号化画像データを生成する画像データ符号化工程のコードと、

前記音声データを複数種類の音声データに階層化して符号化し、複数階層の夫  
々に相当する符号化音声データを生成する音声データ符号化工程のコードと、

前記符号化画像データと前記符号化音声データを用いてフレーム符号化デー  
タを生成するフレーム符号化データ生成工程のコードと、を備え、

前記フレーム符号化データ生成工程においては、所定基準に基づいて決定され  
る同一レベルに属する画像符号化データと音声符号化データをグループ化するこ  
とによってレベルの異なる複数のグループを作成し、該複数のグループを重要度  
の高いレベル順に並べることによって前記フレーム符号化データを生成する  
ことを特徴とするプログラム。

【請求項 2 3】 請求項 2 1 または 2 2 に記載のプログラムを記録した記録  
媒体。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、画像データと音声データを含むフレームデータに対して符号化を行  
う符号化装置及び符号化方法に関するものである。

【0 0 0 2】

【従来の技術】

昨今、パソコンやモバイル端末の普及により、インターネットを介したデジタ  
ルデータの通信（データ通信）が幅広く行われるようになった。データ通信にお

いて流通するデジタルデータのひとつに動画像がある。一般的に動画像はデータ量が大きいため、送信前に動画像中の静止画像とそれに付随する音声単位（フレーム）として符号化されることにより、送信データ量が抑制される。

## 【0003】

このような動画像データの通信方法として、送信データに対し、受信側において動画像データ中の画像データ及び音声データの復号が進むにつれて、復号画像または復号音声の品位が向上するような特性（スケーラビリティ）を持たせる方法が知られている。

## 【0004】

このように送信対象データに対してスケーラビリティを付与する場合、送信対象となる画像データ及び音声データの符号化処理において、スケーラビリティを付与していた。

## 【0005】

## 【発明が解決しようとする課題】

しかしながら、上記従来のスケーラビリティを伴う通信方法においては、送信データに対するスケーラビリティの付与は該データに対する符号化処理の際に行われていたため、既に符号化されている画像データや音声データに対してスケーラビリティを付与したい場合には、これらを一旦復号した後に、改めてスケーラビリティを付与する符号化を行わざるをえなかった。

## 【0006】

また、画像データと音声データの両方に対してスケーラビリティを付与しつつ、これら両方のデータを含む符号化データを生成するような符号化方法は、未だ確立されていない。

## 【0007】

本発明は上述の問題点を鑑みてなされたものであり、既に符号化された画像データ及び音声データの両方に対して、これらを復号することなく適切にスケーラビリティを付与し、これら両方のデータを含む符号化データを生成する符号化装置及び符号化方法を提供することを目的とする。

## 【0008】

また、動画像を構成する各フレームの画像データと音声データを適切な単位でグループ化した符号化データを生成し、伝送することにより、受信側において該符号化データの効率的な利用を可能とする符号化装置及び符号化方法を提供することを目的とする。

【 0 0 0 9 】

【課題を解決するための手段】

本発明の目的を達成するために、例えば本発明の符号化装置は以下の構成を備える。

【 0 0 1 0 】

すなわち、画像データと音声データを含むフレームデータに対して符号化を行う符号化装置であって、前記フレームデータに含まれる前記画像データと前記音声データを分離する分離手段と、分離された前記画像データに対し、その低周波成分から高周波成分へ順次、符号化を行って画像符号化データを生成する画像データ符号化手段と、分離された前記音声データに対し、その低周波成分から高周波成分へ順次、符号化を行って音声符号化データを生成する音声データ符号化手段と、前記画像符号化データと前記音声符号化データを用いてヘッダ情報を生成し、該ヘッダ情報及び前記画像符号化データ及び前記音声符号化データによってフレーム符号化データを生成するフレーム符号化データ生成手段と、を備えることを特徴とする。

【 0 0 1 1 】

また、他の構成によれば、画像データと音声データを含むフレームデータに対して符号化を行う符号化装置であって、前記フレームデータに含まれる前記画像データと前記音声データを分離する分離手段と、前記画像データを複数種類の画像データに階層化して符号化し、複数階層の夫々に相当する符号化画像データを生成する画像データ符号化手段と、前記音声データを複数種類の音声データに階層化して符号化し、複数階層の夫々に相当する符号化音声データを生成する音声データ符号化手段と、前記符号化画像データと前記符号化音声データを用いてフレーム符号化データを生成するフレーム符号化データ生成手段と、を備え、前記フレーム符号化データ生成手段は、所定基準に基づいて決定される同一レベルに

属する画像符号化データと音声符号化データをグループ化することによってレベルの異なる複数のグループを作成し、該複数のグループを重要度の高いレベル順に並べることによって前記フレーム符号化データを生成することを特徴とする。

#### 【 0 0 1 2 】

##### 【発明の実施の形態】

以下に説明する各実施形態において、符号化対象となる動画画は複数のフレームで構成されるものとする。フレームとは、人間が視覚的に動画画を認識する為に順次表示される、1枚毎の静止画像のことであり、各静止画像には、各画像が映し出されている期間（表示期間）に再生される音声が付属している。即ち、1フレームのデータ（フレームデータ）は、1枚の静止画像のデータ（画像データ）と音声のデータ（音声データ）から構成されている。また、フレームデータを符号化してフレーム符号化データを生成するという事は、画像データと音声データのそれぞれを符号化し、画像符号化データと音声符号化データを生成することである。

#### 【 0 0 1 3 】

フレーム符号化データを復号する装置（フレーム復号装置）がフレーム符号化データを復号する際、CPUの能力が足りないこと等の原因により、表示期間内に1つのフレーム符号化データの全てを復号（完全復号）できない場合がある。そのような状況では、画像符号化データと音声符号化データの一部を復号する部分復号が行われる。以下、画像符号化データの部分復号によって得られる画像を部分復号画像、音声符号化データの部分復号によって得られる音声を部分復号音声と称する。同様に、画像符号化データの完全復号によって得られる画像を完全復号画像、音声データの完全復号によって得られる音声を完全復号音声と称する。

#### 【 0 0 1 4 】

画像符号化データの部分復号においては、たとえ低画質であっても完全復号画像の概形が表示される必要がある。

#### 【 0 0 1 5 】

また音声符号化データの部分復号においては、部分復号画像の画質に対応した

程度に、完全復号音声のおおまかな音声が生産されることが望まれる。

【0016】

この完全復号画像の概形の表示と、完全復号音声のおおまかな再生は、画像データと音声データを離散ウェーブレット変換して、それらにスケーラビリティをもたせるように符号化することで達成される。

【0017】

上述したように本発明においては、画像データと音声データの双方がスケーラビリティを持つように符号化し、種々の方法によってフレーム符号化データを生成することを目的とする。

【0018】

以下添付図面に従って、本発明を好適な実施形態に従って説明する。

【0019】

#### <第1実施形態>

##### ●符号化装置構成

図1Aは、本実施形態における符号化装置の構成を示したブロック図である。同図において101はフレームデータ入力部、102はフレームデータ分離部、103は画像データ符号化部、104は音声データ符号化部、105はフレーム符号化データ生成部A、106はフレーム符号化データ出力部である。

【0020】

図1Bは、図1Aにおける画像データ符号化部103の構成を示したブロック図である。同図において107は画像データ入力部、108は離散ウェーブレット変換部A、109はバッファ、110は係数量子化部、111はエントロピー符号化部、112は画像符号化データ生成部A、113は画像符号化データ出力部である。

【0021】

図1Cは、図1Aにおける音声データ符号化部104の構成を示したブロック図である。同図において114は音声データ入力部、115は離散ウェーブレット変換部B、116は音声符号化データ出力部である。

##### ●フレーム符号化処理

以上の構成を有する本実施形態の符号化装置において行われる、フレーム符号化処理のフローチャートを図 1 4 に示し、以下、同図を用いて説明する。

【 0 0 2 2 】

まず図 2 に示されているような、画像データと音声データから構成されるフレームデータが、フレームデータ入力部 1 0 1 に入力され、そしてフレームデータ分離部 1 0 2 に出力される（ステップ S 1 4 0 1）。このフレームデータ入力部 1 0 1 は、例えばデジタルビデオカメラ、デジタルスチルカメラ等の撮像装置、或いは CCD などの撮像デバイス、或いはネットワーク回線のインターフェース等が用いられる。また、フレームデータ入力部 1 0 1 は RAM、ROM、ハードディスク、CD-ROM が用いられる。

【 0 0 2 3 】

なお、符号化対象の動画像中における複数のフレームは 1 つずつ、フレームデータ入力部 1 0 1 に入力されるものとする。またフレームデータ入力部 1 0 1 以降の処理は、フレームデータ毎、独立に行われるものとする。

【 0 0 2 4 】

フレームデータ分離部 1 0 2 に入力されたフレームデータは、図 3 に示されているように、音声データと画像データに分離される（ステップ S 1 4 0 2）。そして画像データは画像データ符号化部 1 0 3、音声データは音声データ符号化部 1 0 4 に入力される。

【 0 0 2 5 】

画像データ符号化部 1 0 3 に入力された画像データは、後述する処理により符号化が行われ、画像符号化データとなる（ステップ S 1 4 0 3）。そして画像符号化データは、フレーム符号化データ生成部 A 1 0 5 に入力される。

【 0 0 2 6 】

音声データ符号化部 1 0 4 に入力された音声データは、後述する処理により符号化が行われ、音声符号化データとなる（ステップ S 1 4 0 4）。そして音声符号化データは、フレーム符号化データ生成部 A 1 0 5 に入力される。

【 0 0 2 7 】

フレーム符号化データ生成部 A 1 0 5 に音声符号化データと画像符号化データ

が入力されると、ヘッダが生成される（ステップ S 1 4 0 5）。なおヘッダには、画像データ符号化部 1 0 3 における画像入力部 1 0 9 に入力された画像のサイズ、画像が 2 値画像であるか多値画像であるかを示すタイプなどの情報、画像符号化データの長さ、音声符号化データの長さ、並びに送信する符号化装置を示す文字列、送信日時、等が書き込まれる。また、画像符号化データの先頭アドレス、音声符号化データの先頭アドレスも書き込まれる。

そして図 4 に示されているように、ヘッダ、音声符号化データ、画像符号化データによりフレーム符号化データが生成される（ステップ S 1 4 0 6）。

#### 【 0 0 2 8 】

フレーム符号化データ出力部 1 0 6 では、入力されたフレーム符号化データが外部へ出力（送信）される（ステップ S 1 4 0 7）。このフレーム符号化データ出力部 1 0 6 には、公衆回線、無線回線、LAN 等のインターフェースを用いることができる。

#### 【 0 0 2 9 】

##### ●画像データ符号化処理

以下に、画像データ符号化部 1 0 3 における画像データの符号化処理（ステップ S 1 4 0 3）について、この処理のフローチャートを図 1 5 に示し、同図を用いて説明する。

#### 【 0 0 3 0 】

本実施形態におけるフレーム中の符号化対象となる画像データは、8 ビットのモノクロ画像データとする。しかしながら、各画素 4 ビット、1 0 ビット、1 2 ビットといった具合に 8 ビット以外のビット数で表すモノクロ画像、或いは各画素における各色成分（RGB / Lab / YCrCb）を 8 ビットで表現するカラーの多値画像データである場合に本実施形態を適用することも可能である。また、画像を構成する各画素の状態等を表す情報が多値情報である場合、例えば各画素の色を表す多値のインデックス値である場合にも適用できる。このような各種多値情報に本実施形態を応用する場合には、これら多値情報を後述するモノクロ画像データに変換すればよい。

#### 【 0 0 3 1 】



まず、画像データ入力部 1 0 7 から符号化対象となる画像データを構成する画素データがラスタースキャン順に入力され、離散ウェーブレット変換部 A 1 0 8 に出力される（ステップ S 1 5 0 1）。

## 【 0 0 3 2 】

離散ウェーブレット変換部 A 1 0 8 は、画像データ入力部 1 0 7 から入力される 1 つの静止画像分の画像データ  $x(n)$  における複数の画素（参照画素）のデータ（参照画素データ）を用いて、離散ウェーブレット変換を行う（ステップ S 1 5 0 2）。

## 【 0 0 3 3 】

以下に、離散ウェーブレット変換後の画像データ（離散ウェーブレット変換係数）を示す。

## 【 0 0 3 4 】

$$r1(n) = \text{floor} \{ (x(2n) + x(2n+1)) / 2 \}$$

$$d1(n) = x(2n+2) - x(2n+3) + \text{floor} \{ (-r1(n) + r1(n+2) + 2) / 4 \}$$

上記変換式において、 $r1(n)$ 、 $d1(n)$  は離散ウェーブレット変換係数列（以下変換係数列）であり、 $r1(n)$  は低周波サブバンド、 $d1(n)$  は高周波サブバンドである。また、上式において  $\text{floor} \{ X \}$  は  $X$  を超えない最大の整数値を表す。この離散ウェーブレット変換を模式的に表わしたのが図 5 である。

## 【 0 0 3 5 】

上記変換式は一次元のデータに対するものであるが、この変換を水平方向、垂直方向の順に適用して二次元の変換を行うことにより、参照画素データを図 6（a）に示すように LL, HL, LH, HH の 4 つのサブバンドに分割することができる。ここで、L は低周波サブバンド、H は高周波サブバンドを示している。次に LL サブバンドを、同様に 4 つのサブバンドに分け（図 6（b））、その中の LL サブバンドをまた 4 サブバンドに分ける（図 6（c））ことによって、合計 1 0 個のサブバンドを作成する。

## 【 0 0 3 6 】

図 6（c）において、各サブバンドの名称における数字をそれぞれのサブバンドのレベルとする。つまり、HL 1, HH 1, LH 1 はレベル 1 のサブバンドで

あり、HL 2, HH 2, LH 2はレベル 2 のサブバンドである。なお LL サブバンドひとつしかないので添字を付けず、レベル 0 のサブバンドとする。

## 【 0 0 3 7 】

また、レベル 0 からレベル n までのサブバンドを復号することで得られる復号画像を、レベル n の復号画像と呼ぶ。復号画像は、そのレベルが高い程、解像度が高い。つまり上述の通りに離散ウェーブレット変換された画像データは、部分復号されることで原画像の概形を表示可能とする。

## 【 0 0 3 8 】

図 6 (c) に示す 10 個のサブバンドは、バッファ 109 に一旦格納され、LL, HL 1, LH 1, HH 1, HL 2, LH 2, HH 2, HL 3, LH 3, HH 3 の順に、つまり、レベルが低いサブバンドからレベルが高いサブバンドの順に、係数量子化部 110 へ出力される。

## 【 0 0 3 9 】

係数量子化部 110 では、バッファ 109 から出力される各サブバンドの変換係数を周波数成分毎に定めた量子化ステップで量子化し、量子化後の値（係数量子化値）をエントロピー符号化部 111 へ出力する（ステップ S 1503）。ここで、変換係数値を X、該係数の属するサブバンド（周波数成分）に対する量子化ステップの値を q とするとき、量子化後の係数値（係数量子化値）Q(X) は次式によって求められる。

## 【 0 0 4 0 】

$$Q(X) = \text{floor} \{ (X/q) + 0.5 \}$$

図 7 に、本実施形態における各周波数成分と量子化ステップとの対応を示す。同図に示す様に、低周波サブバンド（LL 等）よりも高周波サブバンド（HL 3、LH 3、HH 3 等）の方に、大きい量子化ステップが与えられている。一つのサブバンドにおける全ての変換係数を量子化した後、それら係数量子化値 Q(X) をエントロピー符号化部 111 へ出力する。

## 【 0 0 4 1 】

エントロピー符号化部 111 では、入力された係数量子化値を算術符号化によりエントロピー符号化し、エントロピー符号化値を生成する（ステップ S 150

4)。生成されたエントロピー符号化値は、画像符号化データ生成部 A 1 1 2 に出力され、図 8 に示されるように低レベルのサブバンドから高レベルのサブバンドの順にサブバンド単位で並べられることによって、画像符号化データが生成される（ステップ S 1 5 0 5）。

#### 【0042】

上述のように生成された画像符号化データは、画像符号化データ出力部 1 1 3 を介してフレーム符号化データ生成部 A 1 0 5 に出力される。

#### 【0043】

#### ●音声データ符号化処理

次に、音声データ符号化部 1 0 4 における音声データの符号化処理（ステップ S 1 4 0 4）について、同処理のフローチャートを図 1 6 に示し、同図を用いて説明する。

#### 【0044】

本実施形態におけるフレーム中の符号化対象となる音声データは、音声データ入力部 1 1 4 から入力され、離散ウェーブレット変換部 B 1 1 5 に出力される（ステップ S 1 6 0 1）。

#### 【0045】

離散ウェーブレット変換部 B 1 1 5 は、音声データ入力部 1 1 4 から入力される音声データ  $y(n)$  に対して離散ウェーブレット変換を施す（ステップ S 1 6 0 2）。

#### 【0046】

以下に、ウェーブレット変換後の音声データ（離散ウェーブレット変換係数）を示す。

#### 【0047】

$$r2(n) = \text{floor} \{ (y(2n) + y(2n+1)) / 2 \}$$

$$d2(n) = y(2n+2) - y(2n+3) + \text{floor} \{ (-r2(n) + r2(n+2) + 2) / 4 \}$$

上記変換式において、 $r2(n)$ 、 $d2(n)$  はウェーブレット変換係数列であり、 $r2(n)$  は低周波サブバンド、 $d2(n)$  は高周波サブバンドである。

#### 【0048】

本変換方式においてはまず、音声データは図 9 (a) に示すように L, H の 2 つのサブバンドに分割される。ここで、L は低周波サブバンド、H は高周波サブバンドを示している。続いて L サブバンドを同様に 2 つのサブバンドに分け (図 9 (b))、その中の L サブバンドをまた 2 サブバンドに分ける (図 9 (c)) ことによって、合計 4 サブバンドを作成する。図 9 (c) に示すように、4 個のサブバンドそれぞれを L, H 1, H 2, H 3 と呼ぶことにする。ここで、各サブバンドの名称における数字を該サブバンドのレベルとする。つまり、H 1 はレベル 1 のサブバンドであり、H 2 はレベル 2 のサブバンド、H 3 はレベル 3 のサブバンドである。なお、L サブバンドはひとつしかないので添え字をつけず、レベル 0 のサブバンドとする。

## 【 0 0 4 9 】

また、レベル 0 からレベル n までのサブバンドを復号することで得られる復号音声、レベル n の復号音声と呼ぶ。復号音声はそのレベルが高い程、元の音声に近い。つまり上述の通りに離散ウェーブレット変換された音声データは、部分復号されることで、元の音声をおおまかに再生できる。

## 【 0 0 5 0 】

図 9 (c) に示す 4 個のサブバンドは、音声符号化データとして音声符号化データ出力部 1 1 6 を介してフレーム符号化データ生成部 A 1 0 5 に出力される (ステップ S 1 6 0 3)。

## 【 0 0 5 1 】

以上説明したように、本実施形態における符号化装置及び符号化方法によれば、フレームデータ中の画像データと音声データを離散ウェーブレット変換して、それぞれにスケーラビリティ機能を持たせることが可能である。そのためフレーム符号化データの復号時に、表示期間内に 1 つのフレーム符号化データの完全復号が行われず部分復号が行われても、原画像の概形の表示、並びに元の音声の大まかな音声の再生は達成される。

## 【 0 0 5 2 】

なお図 1 4, 1 5, 1 6 に示したフローチャートに従ったプログラムコードは不図示のメモリ (ROM や RAM など) や外部記憶装置などに格納され、不図示

のCPUにより読み出され、実行されるものとする。

【0053】

＜第2実施形態＞

以下、本発明に係る第2実施形態について説明する。

【0054】

一般にサーバ／クライアントモデルにおいて、サーバはクライアントが要求するデータを送信する。この送信において、サーバとクライアントを結ぶ回線のデータ転送能力が異なること等により、各クライアントがサーバに要求するデータ量は異なる。従って、各クライアントが要求するデータ量に対応して、サーバが蓄積するデータは、その一部もしくは全部が取り出されてクライアントに送信される。なおデータの一部が取り出されてクライアントに送信される場合、データの一部といえどもクライアントにとって意味があるデータである必要がある。ここで、サーバがデータの一部を取り出しクライアントに送信することを、データの一部送信と呼ぶことにする。

【0055】

画像データと音声データが離散ウェーブレット変換されて生成されたフレーム符号化データの一部送信の際には、画像符号化データ、音声符号化データ共に、サブバンド単位で送信されることが望まれる。さらに、復号画像の画質と復号音声の音質を対応させるために、一部送信される画像符号化データのサブバンドのレベルと、音声符号化データのサブバンドのレベルは対応していることが好ましい。

【0056】

そこで第2実施形態においては、フレーム符号化データの一部送信を受信して復号画像及び復号音声を再生する際に、再生されたフレーム内での画質及び音質が対応するように、画像符号化データと音声符号化データの各レベルをグループ化してフレーム符号化データを生成することを特徴とする。

【0057】

図10に、第2実施形態における符号化装置のブロック構成を示す。第2実施形態の符号化装置は、第1実施形態の符号化装置を構成するフレーム符号化デ

タ生成部 A 1 0 5 を、フレーム符号化データ生成部 B 1 0 0 1 に置換したものである。なお、他の構成については第 1 実施形態と同様であるため、第 1 実施形態と同一の参照番号を付して説明を省略する。

#### 【 0 0 5 8 】

フレーム符号化データ生成部 B 1 0 0 1 においては、音声符号化データと画像符号化データが入力されると、第 1 の実施形態と同様にヘッダが生成される。そしてヘッダ、音声符号化データ、及び画像符号化データからフレーム符号化データが生成される。その際、図 1 1 に示されるように、画像符号化データと音声符号化データの同レベルのサブバンドがグループ化されて、フレーム符号化データが生成される。

#### 【 0 0 5 9 】

図 1 1 に示すフレーム符号化データにおいて、レベル 0 のデータは、画像符号化データのレベル 0 のサブバンド (L L) と音声符号化データのレベル 0 のサブバンド (L) がグループ化されており、また、レベル 1 のデータは、画像符号化データのレベル 1 のサブバンド (H L 1, L H 1, H H 1) と音声符号化データのレベル 1 のサブバンド (H 1) がグループ化されている。なお、レベル 2, レベル 3 についても同様に、画像符号化データと音声符号化データの同レベルのサブバンドがグループ化される。

#### 【 0 0 6 0 】

なお、第 2 実施形態における符号化処理は、第 1 実施形態に示した図 1 4 のフローチャートにおいて、ステップ S 1 4 0 6 のフレーム符号化データ生成処理の際に、図 1 1 に示されているように、画像符号化データと音声符号化データの同レベルのサブバンドをグループ化する処理を追加したものとなる。また、画像データ符号化部 1 0 3, 音声データ符号化部 1 0 4 における各処理のフローチャートは第 1 の実施形態と同様であり、夫々図 1 5, 1 6 に示したフローチャートに従う。

#### 【 0 0 6 1 】

又、このフローチャートに従ったプログラムコードは不図示のメモリ (RAM や ROM 等) や外部記憶装置などに格納され、不図示の CPU により読み出され

、実行されるものとする。

【 0 0 6 2 】

以上説明したように、第 2 実施形態における符号化装置及び符号化方法によれば、画像符号化データと音声符号化データのサブバンドの各レベルをグループ化してフレーム符号化データを生成することで、フレーム符号化データの一部送信に基づいて復号画像と復号音声を再生する際にも、再生されたフレーム内における画質及び音質を適切に対応させることができる。

【 0 0 6 3 】

又、第 2 実施形態の符号化装置及び符号化方法によれば、第 1 実施形態における符号化装置及び符号化方法と同様の効果をも奏することは、上述の説明により明白である。

【 0 0 6 4 】

### ＜第 3 実施形態＞

以下、本発明に係る第 3 実施形態について説明する。

【 0 0 6 5 】

上述した第 2 実施形態で説明したように、サーバとクライアントを結ぶ回線のデータ転送能力が異なること等により、各クライアントがサーバに要求するデータ量は異なる。

【 0 0 6 6 】

この転送能力が異なる回線の種類が複数ある場合、それぞれの回線の転送能力に対応させてフレーム符号化データを生成することは、例えば一部送信の高速化にとって好ましい。

【 0 0 6 7 】

そこで第 3 実施形態においては、フレーム符号化データの一部送信を行う際に、使用する回線の転送速度に応じてフレーム符号化データを生成することを特徴とする。

【 0 0 6 8 】

第 3 実施形態における符号化装置の構成を図 1 2 に示す。第 3 実施形態における符号化装置は、第 1 の実施形態の符号化装置を構成するフレーム符号化データ

生成部 A 1 0 5 を、フレーム符号化データ生成部 C 1 2 0 1 に置換したものである。

【 0 0 6 9 】

ここで、第 3 実施形態における符号化装置において生成されたフレーム符号化データが蓄積されるサーバには、回線 A と回線 B の 2 種類の回線が接続されており、回線 A はフレーム符号化データの一部しか送信できず、回線 B はフレーム符号化データの全部を十分に送信できるものとする。

【 0 0 7 0 】

フレーム符号化データ生成部 C 1 2 0 1 においては、音声符号化データと画像符号化データが入力されると、第 1 実施形態と同様にヘッダが生成される。そしてヘッダ、音声符号化データ、及び画像符号化データからフレーム符号化データが生成される。

【 0 0 7 1 】

なお、図 1 3 に示されているように、ヘッダを除くフレーム符号化データ（準フレーム符号化データ）は、準フレーム符号化データ 1 と準フレーム符号化データ 2 から構成される。このそれぞれの準フレーム符号化データは、画像符号化データと音声符号化データから構成される。準フレーム符号化データ 1 は、画像符号化データと音声符号化データの低周波成分から、回線 A が転送できる符号量に対応して取り出されたものである。準フレーム符号化データ 2 は、画像符号化データと音声符号化データから準フレーム符号化データ 1 を除いたものである。なお、ここで回線 A、B 夫々の転送できる符号量は予め分かっているものとし、その値は予め所定のメモリ（ROM や RAM など）に格納されているものとする。

【 0 0 7 2 】

上述のようにフレーム符号化データが生成されることで、サーバがフレーム符号化データを送信する際、それぞれの回線が有する最大の転送速度でフレーム符号化データの送信を図ることができる。

【 0 0 7 3 】

なお、第 3 実施形態における符号化処理は、上述した第 1 実施形態に示した図 1 4 のフローチャートにおいて、ステップ S 1 4 0 6 でフレーム符号化データの



生成処理の際に、使用回線に応じた符号量に対応した画像符号化データ、音声符号化データを取り出す処理を追加したものとなる。

【 0 0 7 4 】

また、画像データ符号化部 1 0 3、音声データ符号化部 1 0 4 における各処理は第 1 実施形態と同様であり、夫々図 1 5、1 6 に示したフローチャートに従う。

【 0 0 7 5 】

又、このフローチャートに従ったプログラムコードは不図示のメモリ（RAM や ROM 等）や外部記憶装置などに格納され、不図示の CPU により読み出され、実行されるものとする。

【 0 0 7 6 】

以上説明したように第 3 実施形態における符号化装置及び符号化方法によれば、フレーム符号化データの一部送信を行う際に、使用する回線の転送速度に応じてフレーム符号化データを生成することができる。

【 0 0 7 7 】

なお、第 3 実施形態においては、転送能力が異なる 2 種類の回線がサーバに接続されている場合について説明したが、転送能力が異なる 3 種類以上の回線がサーバに接続されていてももちろん構わない。

【 0 0 7 8 】

#### ＜第 4 実施形態＞

以下、本発明に係る第 4 実施形態について説明する。

【 0 0 7 9 】

上述した各実施形態においては、音声データにおける低周波成分を重要と仮定し、まずこの低周波成分を送信する例について説明した。しかしながら音声データにおいては、特に人間の声のデータ（スピーチデータ）が、重要度の高いデータとして扱われることが多い。

【 0 0 8 0 】

そこで第 4 実施形態においてはこの点に着目し、音声データを最重要データであるスピーチデータと、該スピーチデータ以外の（重要度の低い）非スピーチデ

ータに分離し、これらを別々に符号化することによって、スピーチ符号化データと非スピーチ符号化データを生成することを特徴とする。そしてさらに、画像符号化データにおける重要なデータ（低周波サブバンド）と、上記スピーチ符号化データとを最重要レベルのグループとしてグループ化し、その他の画像及び音声データについてもレベル別にグループ化することによって、フレーム符号化データを生成する。

#### 【0081】

以下、第4実施形態における符号化方法について説明する。

#### 【0082】

図17は、第4実施形態における符号化装置の構成を示すブロック図である。これは、第1実施形態において図14に示した音声データ符号化部104及びフレーム符号化データ生成部A105を、それぞれ音声データ符号化部A1701及びフレーム符号化データ生成部D1702に置き換えたものである。

#### 【0083】

また図18は、音声データ符号化部A1701の構成を示すブロック図である。同図において、1801は音声データ分離部、1802はスピーチデータ符号化部、1803は非スピーチデータ符号化部である。

#### 【0084】

以下、以上の構成を有する第4実施形態の符号化装置におけるフレーム符号化処理について説明する。なお、フレームデータ入力部101、フレームデータ分離部102、画像データ符号化部103における処理は上述した第1実施形態と同様であるため詳細な説明は省略し、主に音声データ符号化部A1701の動作について説明する。

#### 【0085】

音声データ符号化部A1701に入力された音声データは、スピーチデータと非スピーチデータに分離される。この音声データの分離方法については、スピーチに該当する周波数成分を分離抽出する等の周知技術が適用可能であるため、ここでは詳細な説明を省略する。ここで分離されたスピーチデータと非スピーチデータは夫々、スピーチデータ符号化部1802、非スピーチデータ符号化部18

03に入力される。

【0086】

スピーチデータ符号化部1802では、入力されるスピーチデータをHVXC(Harmonic Vector eXcitation Coding)で符号化する。また非スピーチデータ符号化部1803は、非スピーチデータをMP3(MPEG Audio Layer III)で符号化する。このようにして生成されたスピーチ符号化データ及び非スピーチ符号化データは、フレーム符号化データ生成部D1702に出力される。

【0087】

フレーム符号化データ生成部D1702においては、図19に示されるように、画像符号化データのLLサブバンドとスピーチ符号化データがグループ化されて準フレーム符号化データ1となり、またLLサブバンド以外の画像符号化データと非スピーチ符号化データがグループ化されて準フレーム符号化データ2となる。そして、ヘッダと準フレーム符号化データ1、2が統合されることによって、フレーム符号化データが生成される。

【0088】

以上説明したように第4実施形態によれば、音声データのうち重要であると考えられるスピーチデータを優先した伝送／復号を可能とするフレーム符号化データを生成することができる。

【0089】

<第5実施形態>

以下、本発明に係る第5実施形態について説明する。

【0090】

上述した第4実施形態においては、音声データをスピーチデータと非スピーチデータの2種類(2階層)に分離したため、音声符号化データも準フレーム符号化データ1、2の2グループに分離されていた。

【0091】

しかしながら、非スピーチデータを更に種々の基準に基づいて2つ以上に分割することにより、音声データをスピーチデータと非スピーチデータ1、非スピーチデータ2、・・・、非スピーチデータnからなる多階層に分離することも考えら

れる。これにより、上述した第2実施形態と同様に、画像と音声とを多階層のグループで構成することができる。

#### 【0092】

第5実施形態においては、音声データを2以上の多階層に分離し、画像データを含めた多階層グループとして符号化する例について説明する。

#### 【0093】

図20は、第5実施形態における符号化装置の構成を示すブロック図である。これは、上述した第1実施形態の図14に示す音声データ符号化部104及びフレーム符号化データ生成部A105をそれぞれ、音声データ符号化部B2001及びフレーム符号化データ生成部E2002に置き換えたものである。

#### 【0094】

また図21は、音声データ符号化部B2001の構成を示すブロック図である。2101は音声データ分離部、2102はスピーチデータ符号化部A、2103は非スピーチデータ符号化部Aである。

#### 【0095】

スピーチデータ符号化部1802においては、スピーチデータを例えばCELP(Code Excited Linear Prediction)で符号化する。また非スピーチデータを、第1階層のモノラル音源と第2階層のステレオ音源に分離し、その上で第1階層をTwin VQ(Transform domain Weighted Interleave Vector Quantization)で符号化し、第2階層をAAC(Advance Audio Codign)で符号化する。ここで符号化された第1階層及び第2階層の非スピーチデータをそれぞれ、第1非スピーチ符号化データ及び第2非スピーチ符号化データと呼ぶ。これらスピーチ符号化データと、第1及び第2非スピーチ符号化データは、フレーム符号化データ生成部E2002に出力される。

#### 【0096】

フレーム符号化データ生成部E2002では、図22に示されるように、画像符号化データのLLサブバンドとスピーチ符号化データがグループ化されて準フレーム符号化データ1となり、HL1,HH1,LH1サブバンドと第1非スピーチ符号化データがグループ化されて準フレーム符号化データ2となり、さらにLL,HL1,HH

1, LH1サブバンド以外のサブバンドと第2非スピーチ符号化データがグループ化されて準フレーム符号化データ3となる。その後、ヘッダと準フレーム符号化データ1, 2, 3が統合されて、フレーム符号化データが生成される。

【0097】

以上説明したように第5実施形態によれば、音声データを多階層のデータに分離し、画像と音声のグループを2つ以上生成することによって、階層的伝送／復号を多段階に行うことが可能となる。

【0098】

なお第5実施形態においては、非スピーチデータを単純に2階層（モノラル音源とステレオ音源）に分離する例について説明したが、本発明はこの例に限定されない。例えば、非スピーチデータに対して離散ウェーブレット変換を施す等によって2つ以上の周波数帯域に分割し、これらを多階層の非スピーチデータの夫々としても構わない。

【0099】

#### <第6実施形態>

以下、本発明に係る第6実施形態について説明する。

【0100】

上述した第2乃至第5実施形態においては、画像符号化データと音声符号化データをグループ化する例をそれぞれ説明した。しかしながら、符号化対象となる動画像（及びこれに付属する音声）の種類によっては、優先すべきデータ種別が異なることが考えられる。例えば、音楽プロモーションビデオ等であれば、高音質な音声データが伝送、復号されることが重要視されるし、一方、スポーツ中継ビデオ等であれば高品位な画像が伝送、復号されることが重要視されるであろう。

【0101】

そこで第6実施形態においては、種々の状況に応じて、符号化データのグループ化方法を選択可能とすることを特徴とする。

【0102】

図23は、第6実施形態における符号化装置の構成を示すブロック図である。

これは、上述した第5実施形態の図20に示すフレーム符号化データ生成部E2002を、フレーム符号化データ生成部F2301に置き換え、さらにグルーピング制御部2302を付加したものである。

【0103】

なお、第6実施形態において生成される画像符号化データ及び音声符号化データは、上述した第5実施形態と同様の多階層形態をなすものとする。

【0104】

画像符号化データと音声符号化データがフレーム符号化データ生成部F2301に入力されると、グルーピング制御部2302が動作し、フレーム符号化データ生成部F2301に対してグループ化の方法に関する指示（グループ化方法指示）を与える。

【0105】

グルーピング制御部2302におけるグループ化方法指示を入力する一形態として、オペレータによる手動入力と考えられる。また別の形態として、グルーピング制御部2302に組み込まれているプログラムによる自動入力がある。なお第6実施形態においては、選択可能なグループ化方法として、“ノーマル”，“画質優先”，“音質優先”の3種類を想定する。

【0106】

グルーピング制御部2302からのグループ化方法指示を受けたフレーム符号化データ生成部F2302は、該指示に基づいて符号化データを生成する。例えば、グループ化方法指示が“ノーマル”であれば、第5実施形態と同様に多階層のグループ化が行われる。また、グループ化方法指示が“画質優先”であれば、図24に示すようにレベル0及びレベル1の画像データを優先して第1グループ（準フレーム符号化データ1）とするようなグループ化が行われる。また、グループ化方法指示が“音声優先”であれば、図25に示すようにレベル0の画像データと全階層の音声データをまとめて第1グループ（準フレーム符号化データ1）とするようなグループ化が行われる。

【0107】

以上説明したように第6実施形態によれば、様々なグループ化を選択的に行う

ことが可能となる。

【0108】

なお第6実施形態においては、説明を簡単するためにグループ化の種類を3種類としたが、もちろんグループ化方法の種類は上記3種類に限定されず、例えば”画質優先”や”音質優先”におけるグループ化の方法として、さらに複数種類を備えることも可能である。

【0109】

また、上述した様な”画質優先”や”音質優先”という概念に基づいてグループ化方法を選択する場合に限らず、様々な状況に適応させる為に、上述した各実施形態に示したグループ化方法を、1つの装置において選択的に使用できる構成を備える場合も、本発明の範疇に含まれる。

【0110】

<第7実施形態>

以下、本発明に係る第7実施形態について説明する。

【0111】

上述した各実施形態において生成されたフレーム符号化データを伝送する場合、回線状況の変動や復号側のCPUパワー等に応じて、低ビットレートの伝送用に割り当てられるフレーム符号化データを変動させる必要が生じてくる。

【0112】

そこで第7実施形態においては、復号側の状況に応じてグループ化方法を適応的に切り替え可能とすることを特徴とする。

【0113】

図26は、第7実施形態における符号化装置の構成を示すブロック図である。これは、上述した第6実施形態の図23に示すフレーム符号化データ生成部F2301及びグルーピング制御部2302を、フレーム符号化データ生成部G2601及びグルーピング制御部A2602に置換したものである。

【0114】

なお、第7実施形態において生成される画像符号化データ及び音声符号化データは、上述した第5実施形態と同様の多階層形態をなすものとする。

## 【 0 1 1 5 】

グルーピング制御部 A 2 6 0 2 は、自装置より送信したフレーム符号化データの、復号側装置における復号状況（各フレーム符号化データについて所定時間内にどの程度復号できているか）を示す情報を受信することができる。そして、この復号状況情報を受信すると、現在符号化を行っている最中のフレーム、又はこれから符号化を開始するフレームに対する適切なグループ化方法を決定する。

## 【 0 1 1 6 】

例えば、グルーピング制御部 A 2 6 0 2 において、送信した各フレーム符号化データが受信側の復号器において確実に復号再生できていないという状況が認識された場合には、最低レベルに相当するグループに含まれる画像かつ／又は音声のデータ量を減少させる様なグループ化方法に切り替える。一方、送信した各フレーム符号化データが受信側の復号器において復号再生され、かつ復号時間的にも未だ余裕が有るという状況が認識された場合には、最低レベルに相当するグループに含まれる画像かつ／又は音声のデータ量を増加させる様なグループ化方法に切り替える。

## 【 0 1 1 7 】

フレーム符号化データ生成部 G 2 6 0 1 は、以上の様にグルーピング制御部 A 2 6 0 2 の指示に従って適切なグループ化を行い、フレーム符号化データを生成する。

## 【 0 1 1 8 】

以上説明したように第 7 実施形態によれば、送信した符号化データの復号状況を考慮した、最適なグループ化を行うことができる。

## 【 0 1 1 9 】

## &lt;その他の実施形態&gt;

上述した第 1 乃至 3 実施形態においては、画像データに対する離散ウェーブレット変換と音声データに対する離散ウェーブレット変換を同様の演算方法によって行う例を示したが、この演算方法は互いに異なっても構わない。

## 【 0 1 2 0 】

また、画像符号化データと同様に、離散ウェーブレット変換された音声データ



を量子化しても構わないし、さらに、該量子化された音声データに対して算術符号等のエントロピー符号化を施しても構わない。

【 0 1 2 1 】

フレーム符号化データ中の任意のアドレスにアクセスすることを容易にするため、画像符号化データや音声符号化データの先頭、終端、並びに画像符号化データや音声符号化データにおけるサブバンドの先頭、終端を示すビットをデータ中に付加しても構わない。

【 0 1 2 2 】

また、スピーチデータの符号化方法は、第 4 乃至第 7 実施形態において説明した方法に限定されず、例えば、G.729, G.723.1等を用いることも可能である。また、非スピーチデータの符号化方法としても例えば、HILIN(Harmonic and Individual Lines plus Noise)やBSAC(Bit Slice Arithmetic Coding)等を用いることが可能である。

【 0 1 2 3 】

(変形例)

なお、本発明は複数の機器（例えばホストコンピュータ、インターフェース機器、リーダ、プリンタ等）から構成されるシステムの一部として適用しても、1つの機器（例えばデジタルビデオカメラ、デジタルスチルカメラ等）からなる装置の1部に適用しても良い。

【 0 1 2 4 】

また、本発明は上記実施の形態を実現するための装置及び方法のみに限定されるものではなく、上記システム又は装置内のコンピュータ（CPU或いはMPU）に、上記実施の形態を実現するためのソフトウェアのプログラムコードを供給し、このプログラムコードに従って上記システム或いは装置のコンピュータが上記各種デバイスを動作させることにより上記実施の形態を実現する場合も本発明の範疇に含まれる。

【 0 1 2 5 】

またこの場合、前記ソフトウェアに関するプログラムコード自体が上記実施の形態の機能を実現することになり、そのプログラムコード自体、及びそのプログ

ラムコードをコンピュータに供給するための手段、具体的には上記プログラムコードを格納した記憶媒体は本発明の範疇に含まれる。

【0126】

この様なプログラムコードを格納する記憶媒体としては、例えばフロッピーディスク、ハードディスク、光ディスク、光磁気ディスク、CD-ROM、磁気テープ、不揮発性のメモリカード、ROM等を用いることができる。

【0127】

また、上記コンピュータが、供給されたプログラムコードのみに従って各種デバイスを制御することにより、上記実施の形態の機能が実現される場合だけではなく、上記プログラムコードがコンピュータ上で稼動しているOS（オペレーティングシステム）、或いは他のアプリケーションソフト等と共同して上記実施の形態が実現される場合にもかかるプログラムコードは本発明の範疇に含まれる。

【0128】

更に、この供給されたプログラムコードが、コンピュータの機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに格納された後、そのプログラムコードの指示に基づいてその機能拡張ボードや機能拡張ユニットに備わるCPU等が実際の処理の一部または全部を行い、その処理によって上記実施の形態が実現される場合も本発明の範疇に含まれる。

【0129】

【発明の効果】

以上説明したように本発明によれば、既に符号化された画像データ及び音声データの両方に対して、これらを復号することなく適切にスケーラビリティを付与し、これら両方のデータを含む符号化データを生成することができる。

【0130】

また、動画像を構成する各フレームの画像データと音声データを適切な単位でグループ化した符号化データを生成し、伝送することにより、受信側において該符号化データの効率的な利用が可能となる。

【図面の簡単な説明】

【図1A】

本発明に係る第 1 実施形態における符号化装置の構成を示したブロック図である。

【図 1 B】

画像データ符号化部 1 0 3 の構成を示したブロック図である。

【図 1 C】

音声データ符号化部 1 0 4 の構成を示したブロック図である。

【図 2】

フレームデータの構成を示す図である。

【図 3】

フレームデータの画像データと音声データへの分離を説明する図である。

【図 4】

フレーム符号化データの構成を示す図である。

【図 5】

離散ウェーブレット変換を模式的に表した図である。

【図 6】

離散ウェーブレット変換により生成される各サブバンドを示す図である。

【図 7】

第 1 実施形態における各周波数成分と量子化ステップとの対応を示す図である。

【図 8】

レベルの低い順にサブバンド単位で並べられた画像符号化データを示す図である。

【図 9】

音声データを複数のサブバンドに分割した図である。

【図 1 0】

第 2 実施形態における符号化装置の構成を示すブロック図である。

【図 1 1】

第 2 実施形態におけるフレーム符号化データの構成を示す図である。

【図 1 2】

第 3 実施形態における符号化装置の構成を示すブロック図である。

【図 1 3】

第 3 実施形態におけるフレーム符号化データの構成を示す図である。

【図 1 4】

第 1 実施形態におけるフレーム符号化処理を示すフローチャートである。

【図 1 5】

画像データ符号化処理を示すフローチャートである。

【図 1 6】

音声データ符号化処理を示すフローチャートである。

【図 1 7】

第 4 実施形態における符号化装置の構成を示すブロック図である。

【図 1 8】

音声データ符号化部 A 1 7 0 1 の構成を示すブロック図である。

【図 1 9】

第 4 実施形態におけるフレーム符号化データの構成を示す図である。

【図 2 0】

第 5 実施形態における符号化装置の構成を示すブロック図である。

【図 2 1】

音声データ符号化部 B 2 0 0 1 の構成を示すブロック図である。

【図 2 2】

第 5 実施形態におけるフレーム符号化データの構成を示す図である。

【図 2 3】

第 6 実施形態における符号化装置の構成を示すブロック図である。

【図 2 4】

第 6 実施形態における画質優先時のフレーム符号化データの構成を示す図である。

【図 2 5】

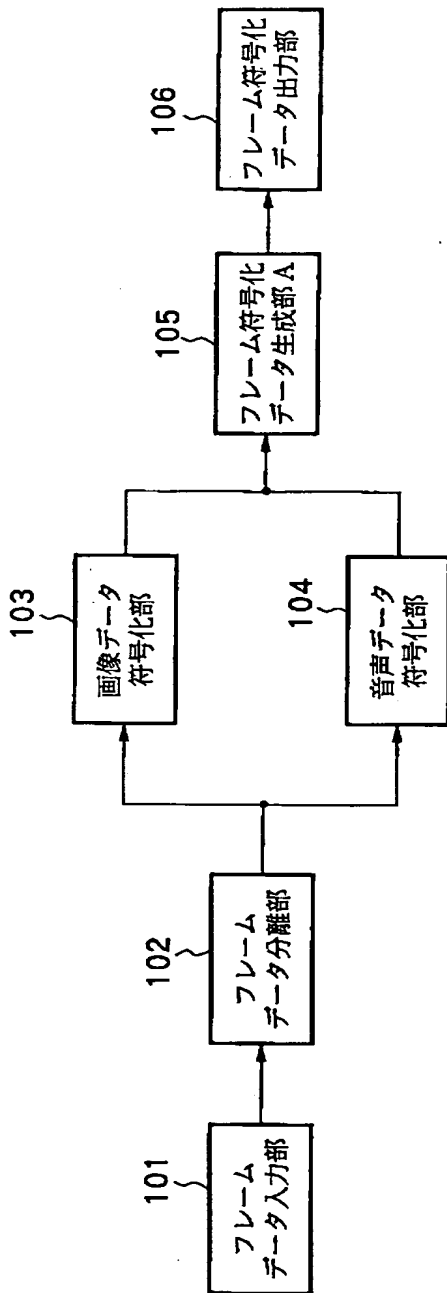
第 6 実施形態における音質優先時のフレーム符号化データの構成を示す図である。

【図 2 6】

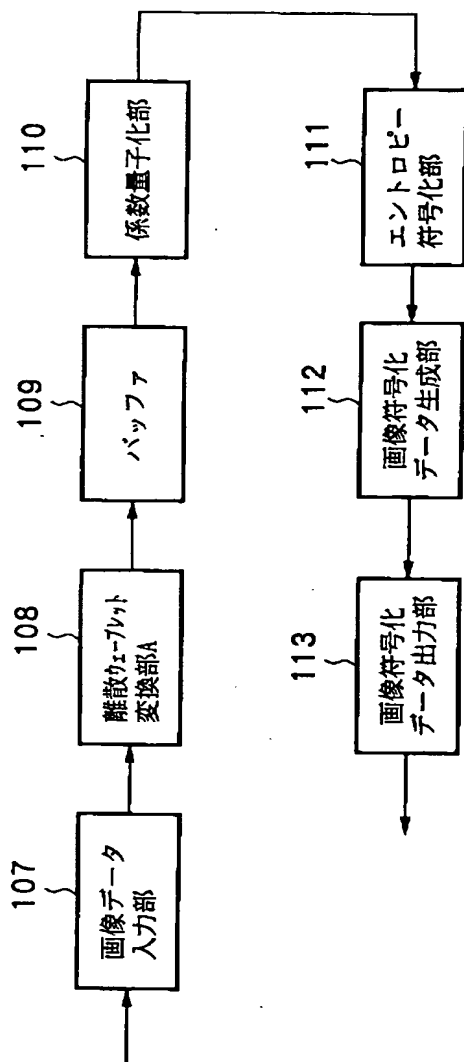
第 7 実施形態における符号化装置の構成を示すブロック図である。

【書類名】 図面

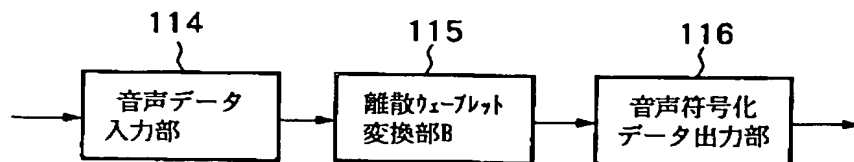
【図 1 A】



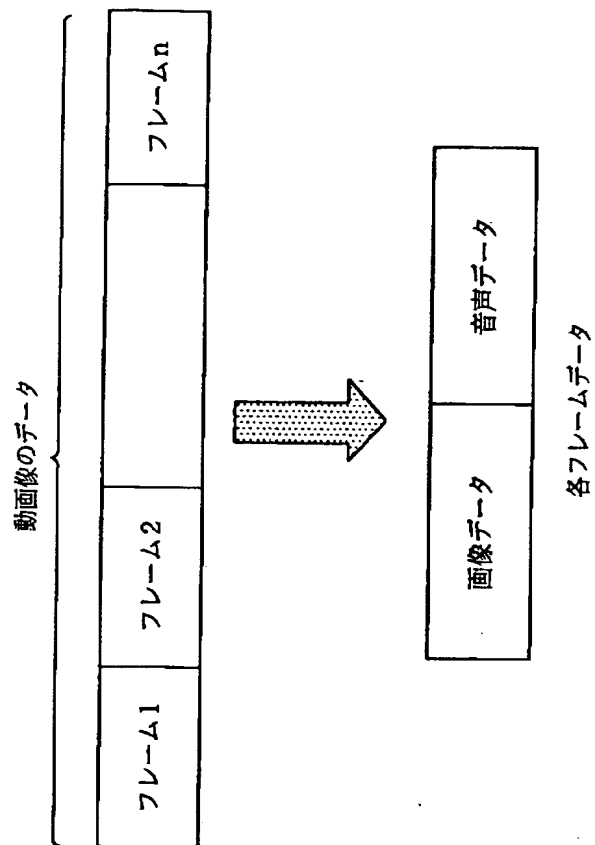
【図 1 B】



【図 1 C】

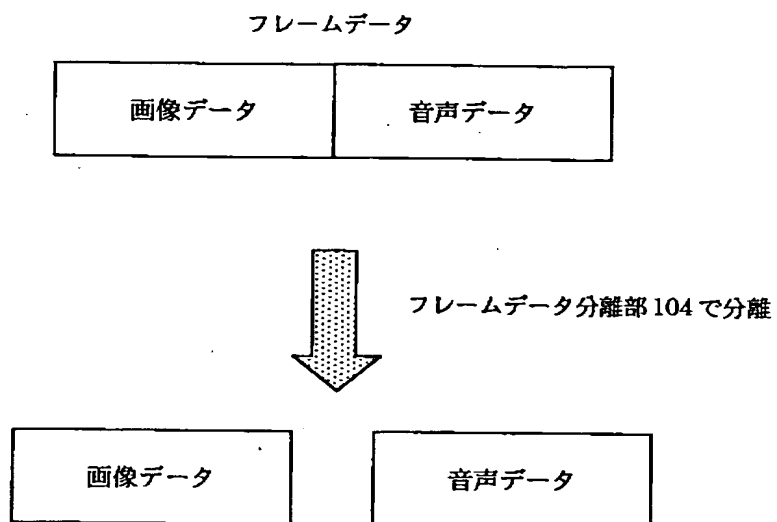


【図 2】

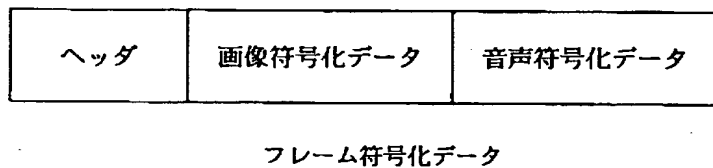




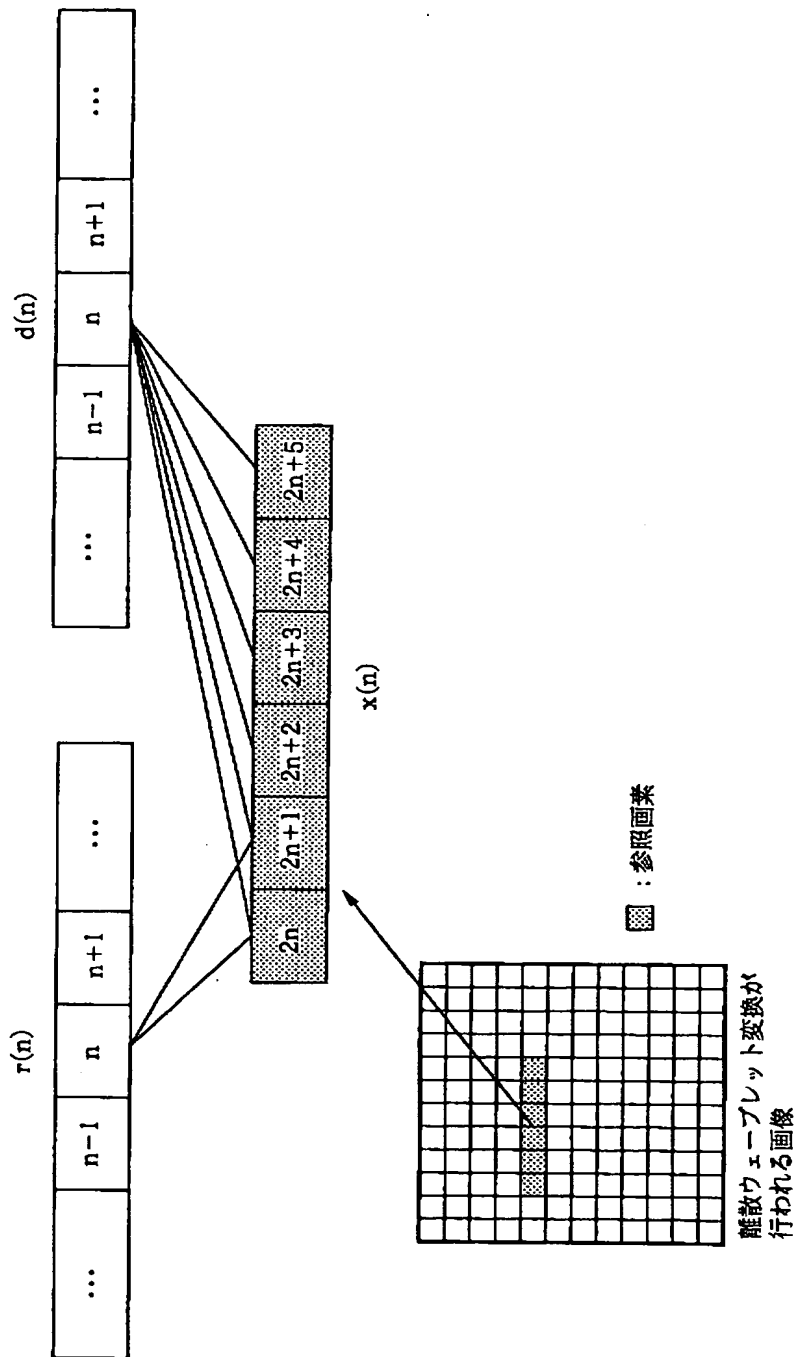
【図 3】



【図 4】



【図5】



【図 6】

LL		HL	
LH		HH	

LL	HL1	HL2		HL3
	LH1	HH1	HH2	
LH2		LH3		HH3

(a)

(b)

(c)

レベル0: LL, レベル1: HL1, HH1, LH1  
 レベル2: HL2, HH2, LH2, レベル3: HL3, HH3, LH3

【図 7】

周波数成分	量子化ステップ
LL	1
HL1	2
HH1	2
LH1	2
HL2	4
HH2	4
LH2	4
HL3	8
HH3	8
LH3	8

【図 8】

画像符号化データ

LL	HL1	LH1	HH1	...	LH3	HH3
----	-----	-----	-----	-----	-----	-----

【図 9】

L	H
---	---

(a)

L	H	H
---	---	---

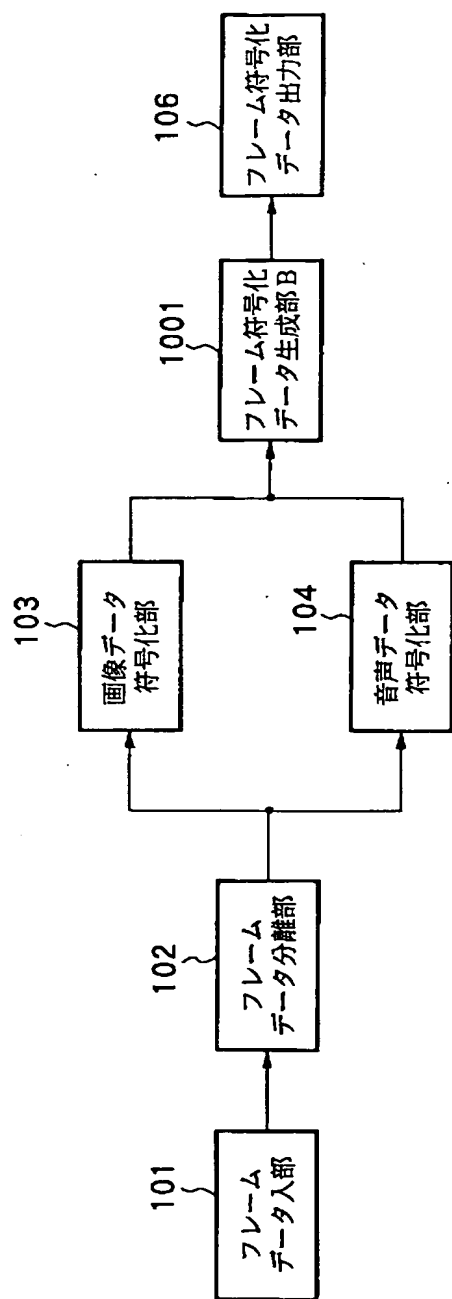
(b)

L	H1	H2	H3
---	----	----	----

(c)

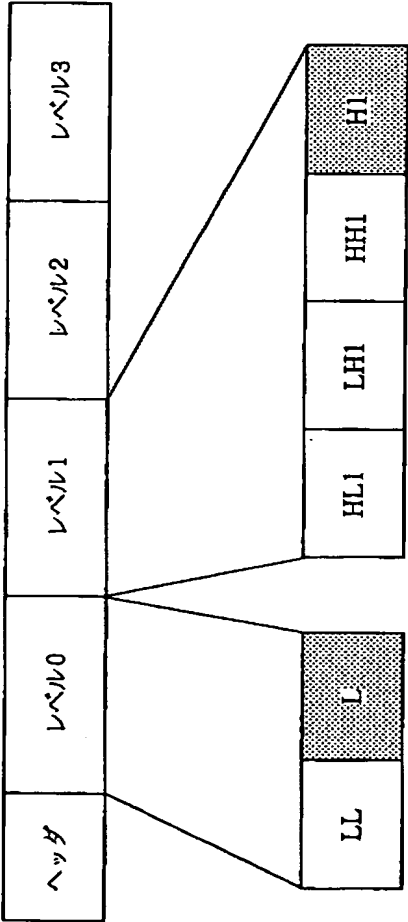
レベル0:L, レベル1:H1, レベル2:H2 レベル3:H3

【図10】



【図 1 1】

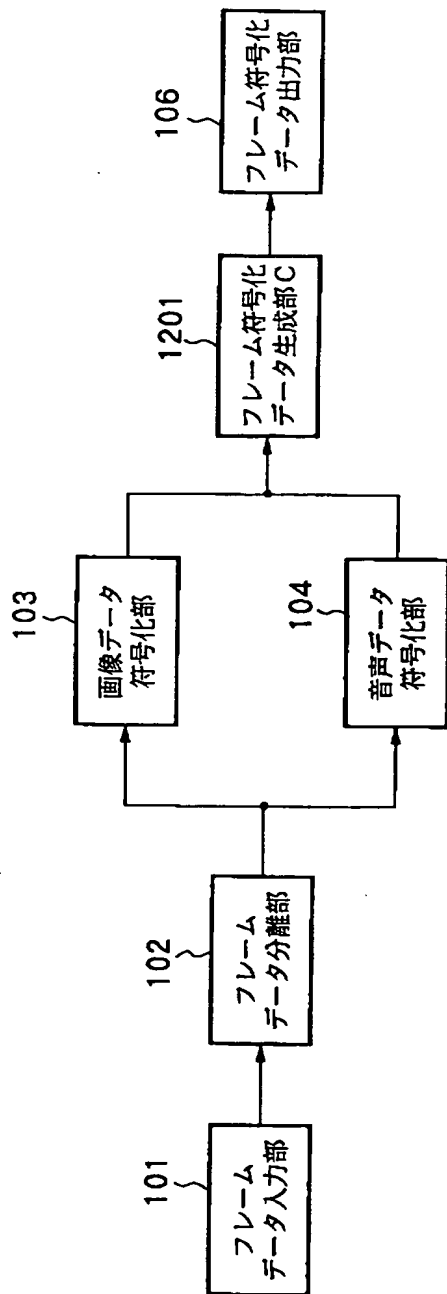
フレーム符号化データ



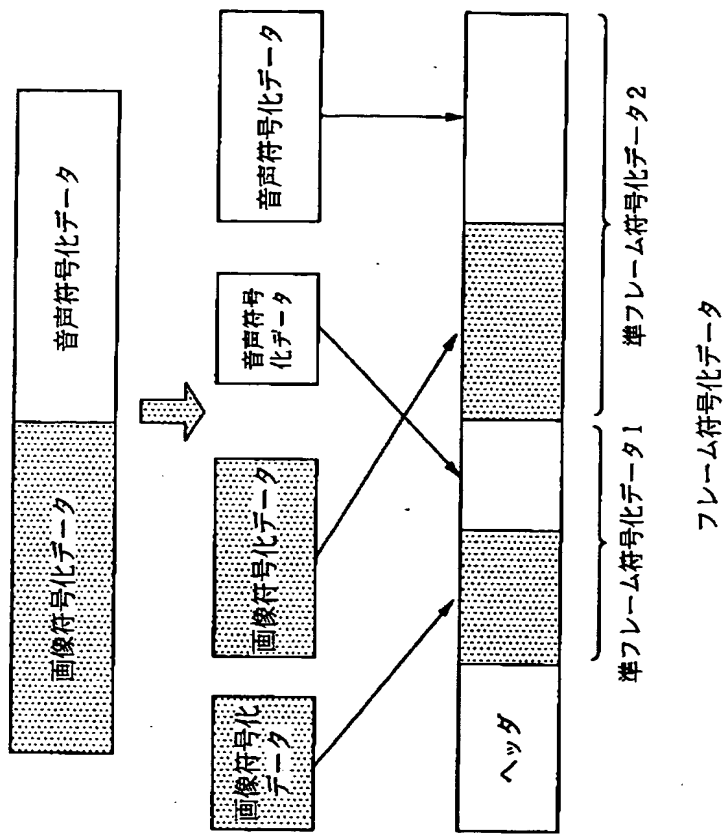
白色のサブバンド: 画像符号化データのサブバンド  
灰色のサブバンド: 音声符号化データのサブバンド  
レベル2, 3の構成は、レベル1 と同様



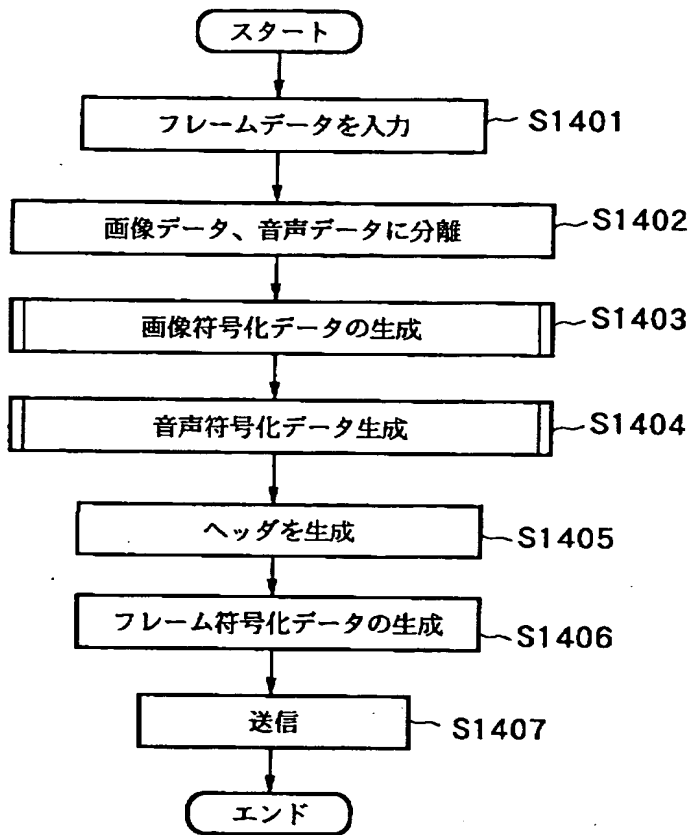
【図 1 2】



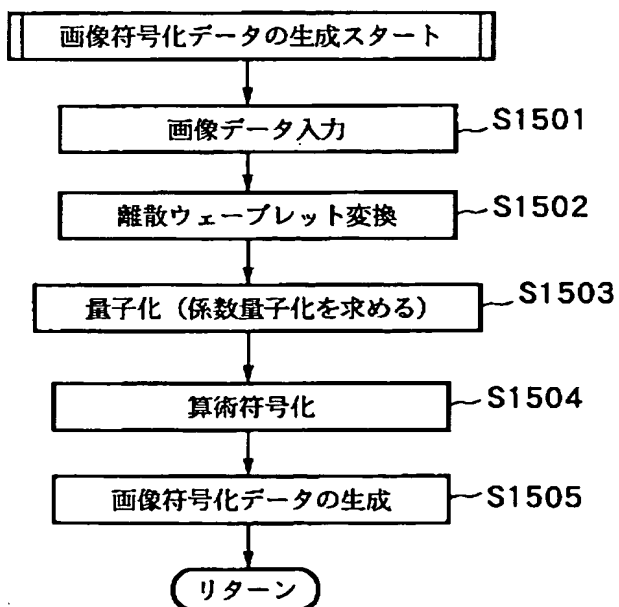
【図13】



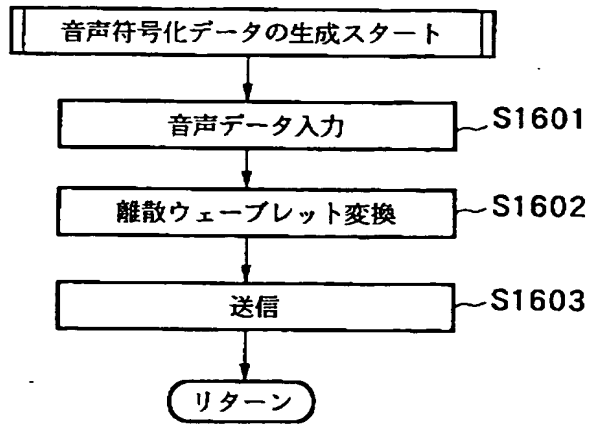
【図 14】



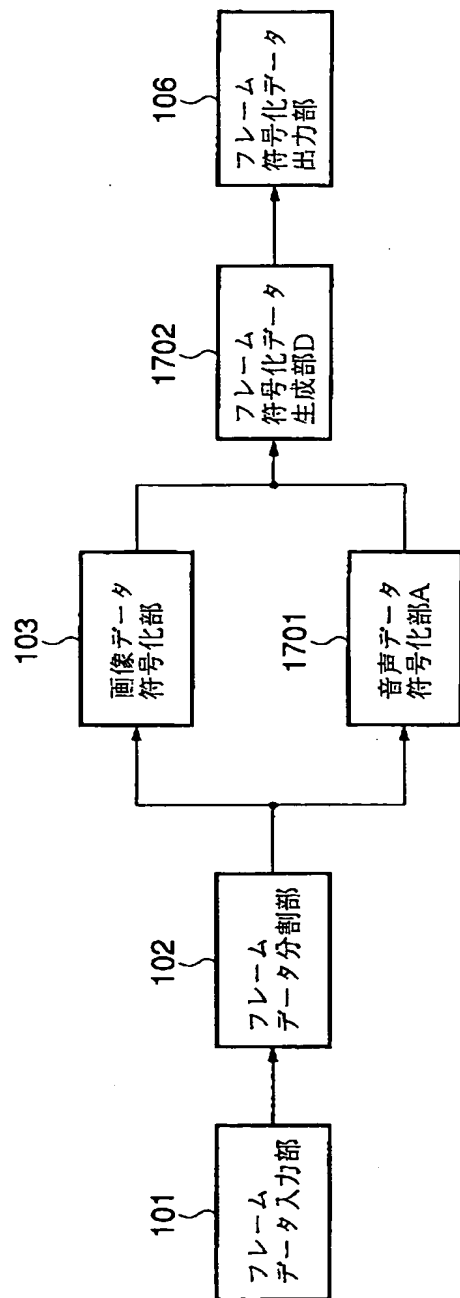
【図 15】



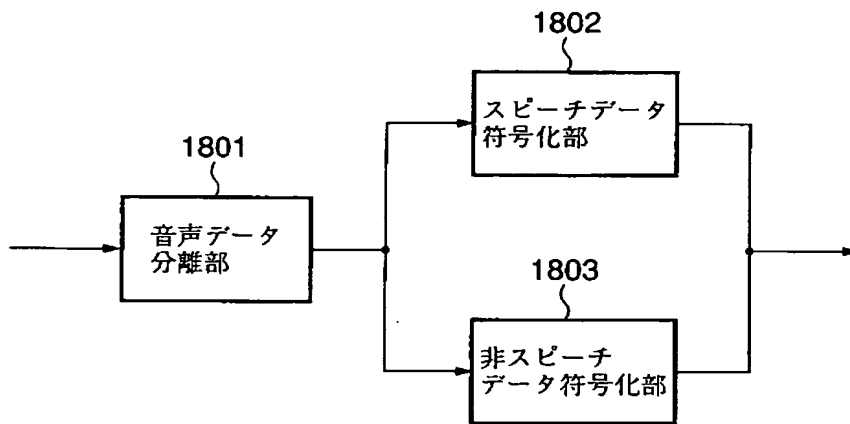
【図 1 6】



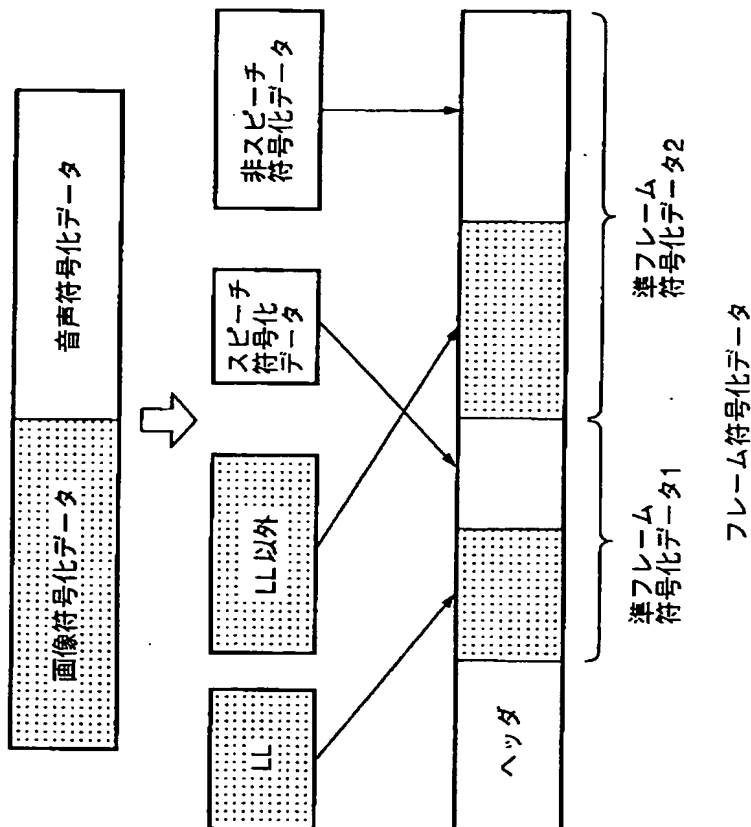
【図 1 7】



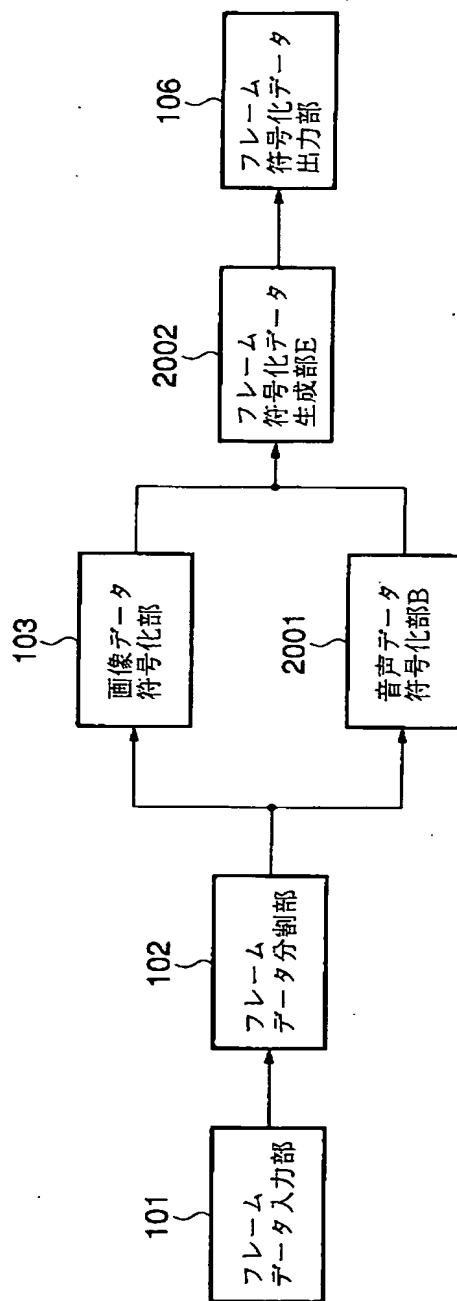
【図 1 8】



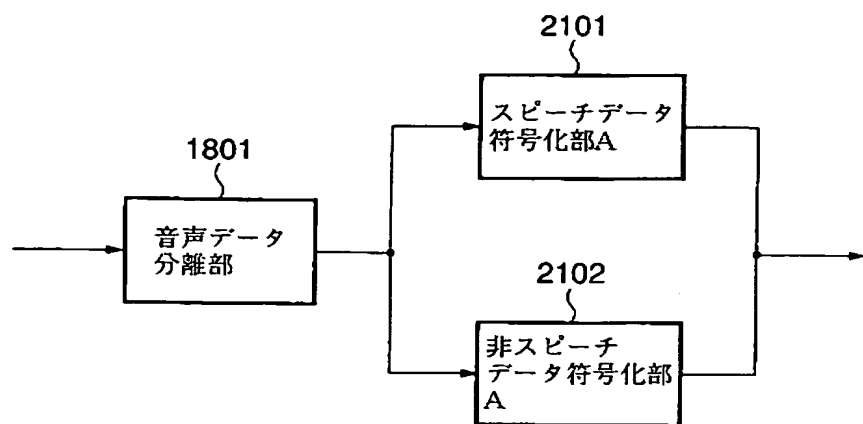
【図 1 9】



【図 2 0】

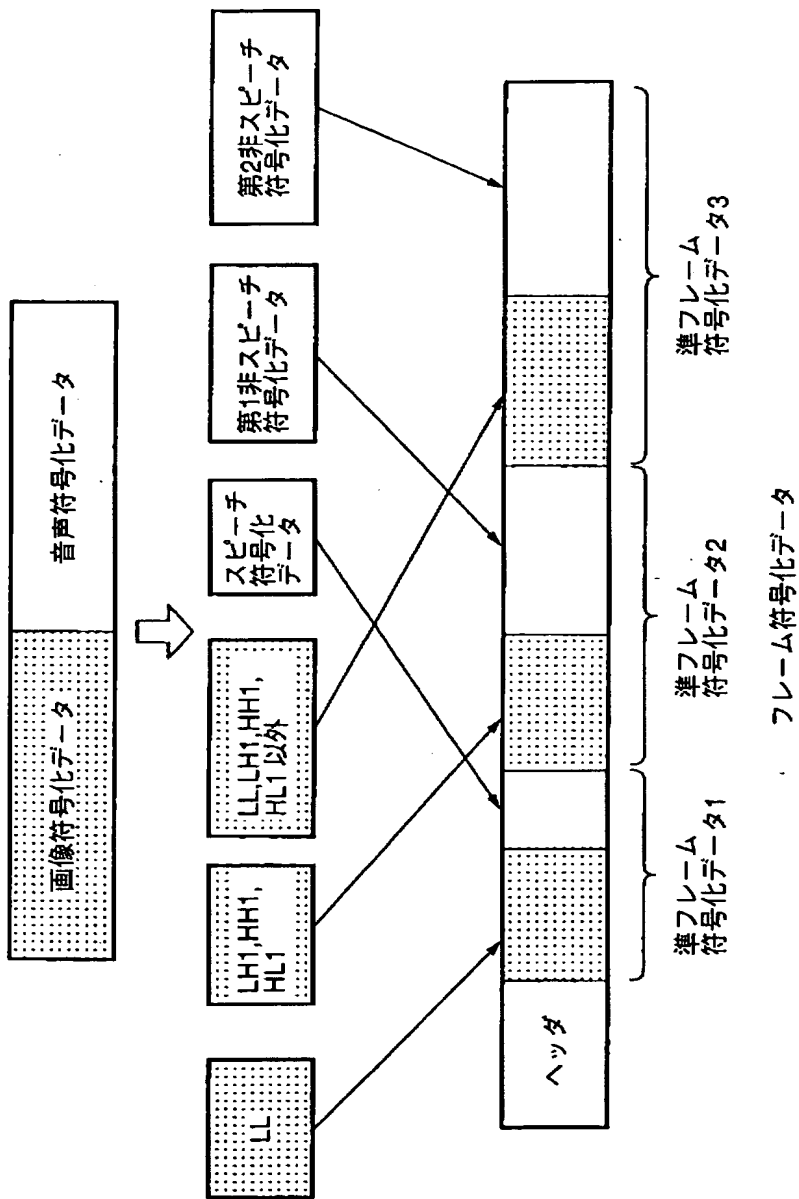


【図 2 1】

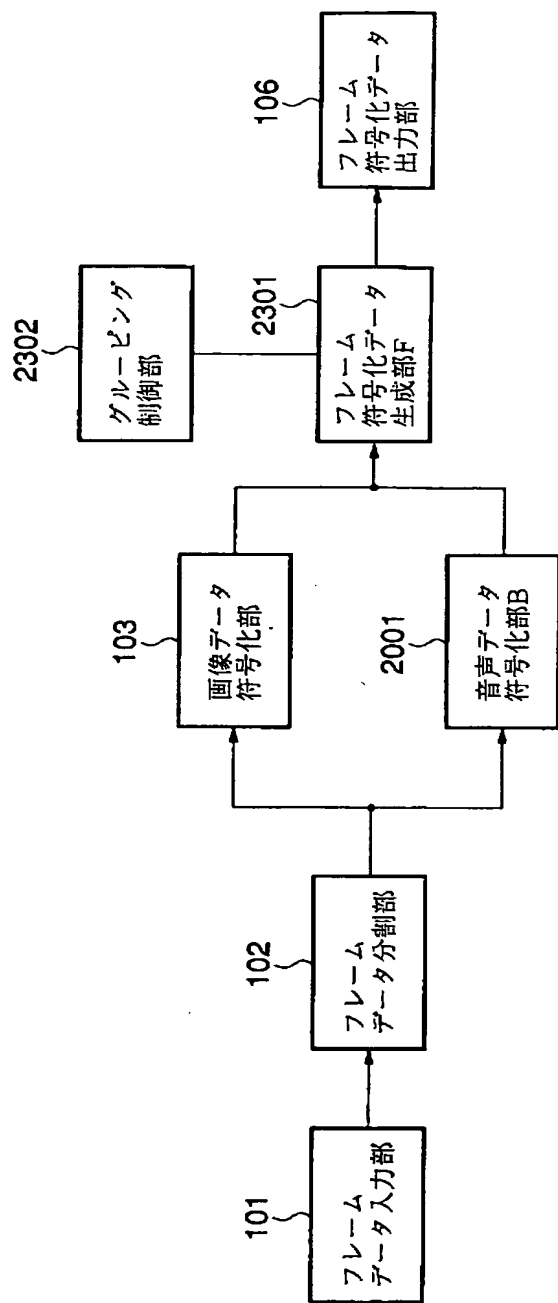




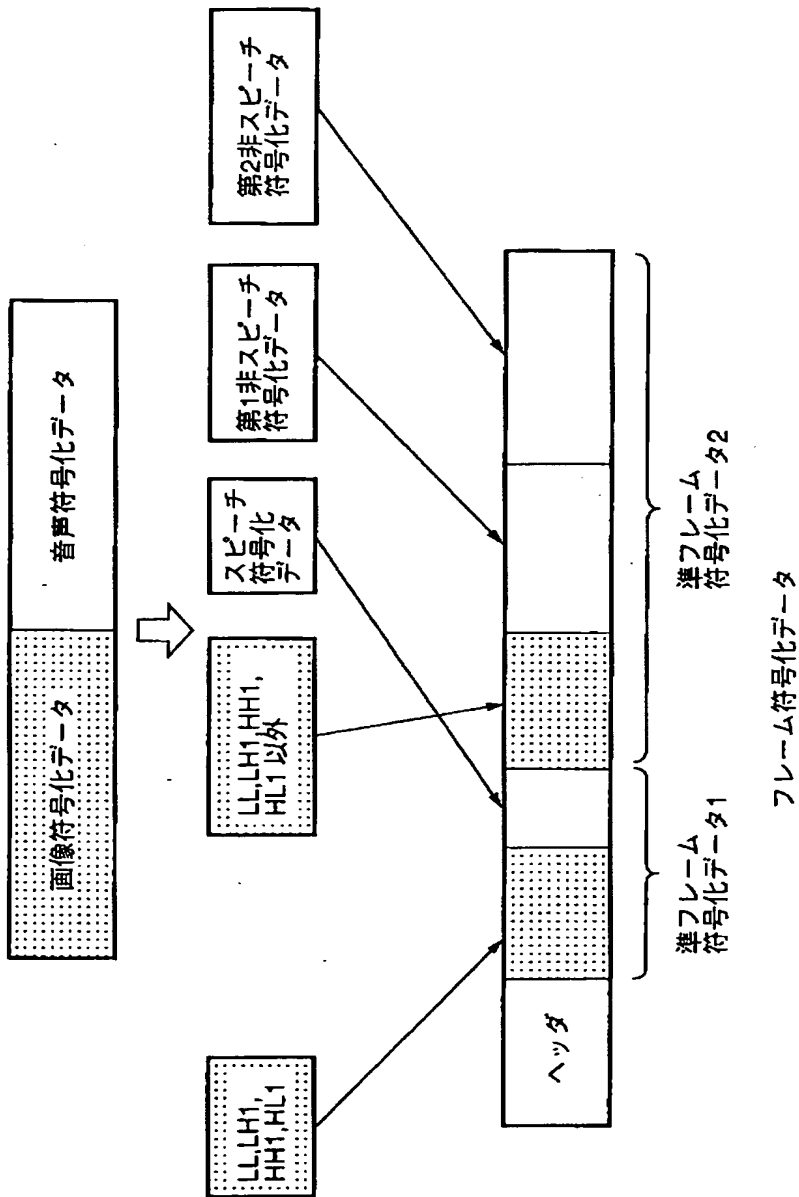
【図 22】



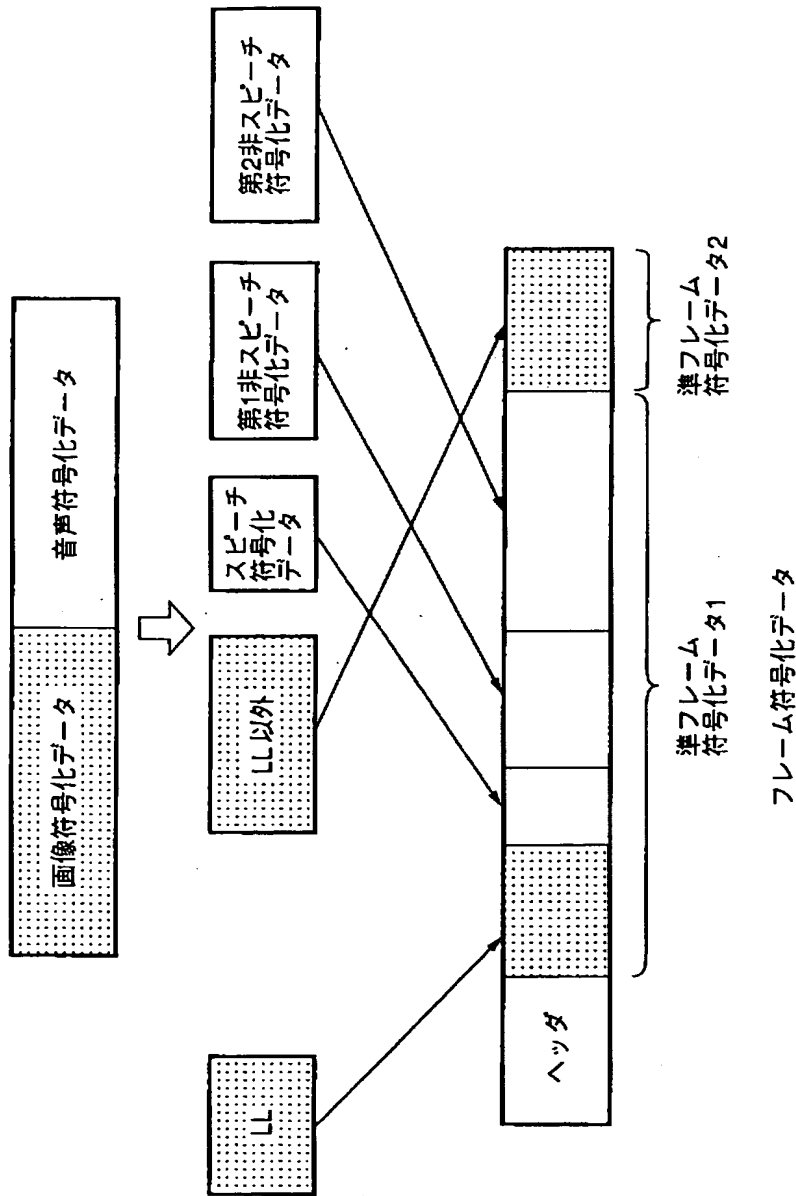
【図 2 3】



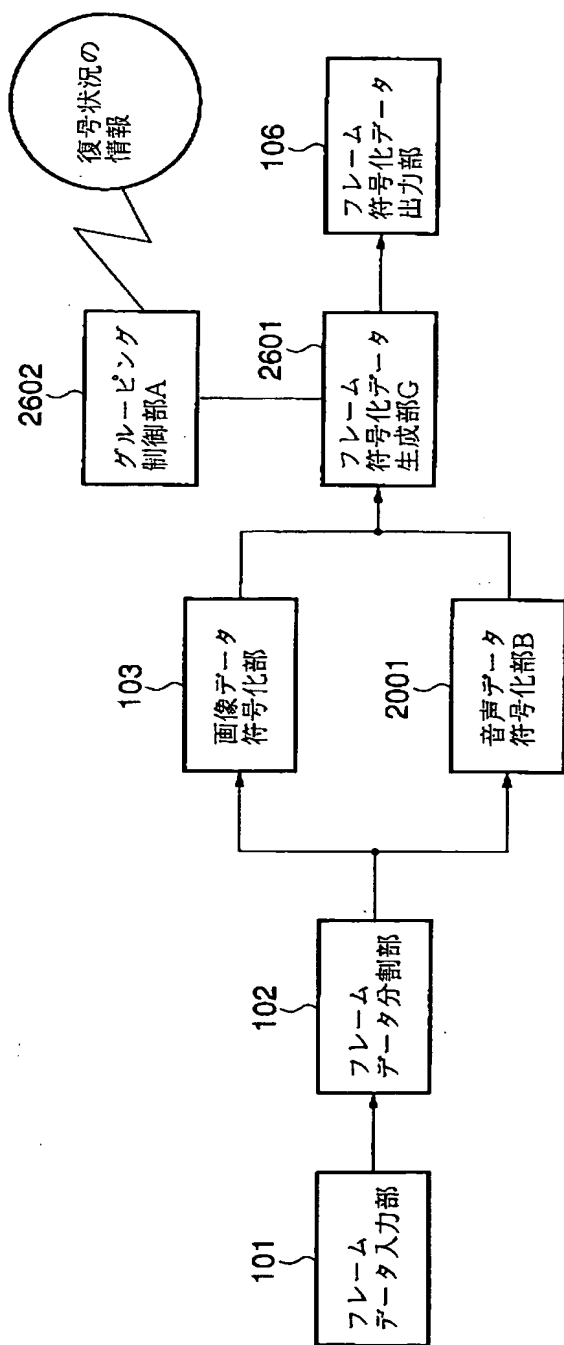
【図 24】



【図 25】



【図 26】



【書類名】 要約書

【要約】

【課題】 既に符号化された画像データ及び音声データの両方に対して、復号することなくスケーラビリティを付与する。また、受信側において該符号化データの効率的な利用を可能とする。

【解決手段】 動画像のフレームデータに含まれる画像データと音声データを分離し、この両方を階層符号化したデータに基づきフレーム符号化データを生成する際に、階層符号化された画像データと音声データにおける重要度が同レベルの周波数サブバンド同士をグループ化し、重要度の高いレベルのグループから順に並べてフレーム符号化データを生成する。

【選択図】 図 1 1

# 認定・付加情報

特許出願の番号	特願 2001-118348
受付番号	50100563561
書類名	特許願
担当官	末武 実 1912
作成日	平成 13 年 4 月 26 日

## <認定情報・付加情報>

### 【特許出願人】

【識別番号】	000001007
【住所又は居所】	東京都大田区下丸子3丁目30番2号
【氏名又は名称】	キャノン株式会社

### 【代理人】

申請人	
【識別番号】	100076428
【住所又は居所】	東京都千代田区紀尾井町3番6号 秀和紀尾井町 パークビル7F 大塚国際特許事務所
【氏名又は名称】	大塚 康德

### 【選任した代理人】

【識別番号】	100112508
【住所又は居所】	東京都千代田区紀尾井町3番6号 秀和紀尾井町 パークビル7F 大塚国際特許事務所
【氏名又は名称】	高柳 司郎

### 【選任した代理人】

【識別番号】	100115071
【住所又は居所】	東京都千代田区紀尾井町3番6号 秀和紀尾井町 パークビル7F 大塚国際特許事務所
【氏名又は名称】	大塚 康弘

### 【選任した代理人】

【識別番号】	100116894
【住所又は居所】	東京都千代田区紀尾井町3番6号 秀和紀尾井町 パークビル7F 大塚国際特許事務所
【氏名又は名称】	木村 秀二

出 願 人 履 歴 情 報

識別番号 [ 0 0 0 0 0 1 0 0 7 ]

1. 変更年月日 1 9 9 0 年 8 月 3 0 日

[ 変更理由 ] 新規登録

住 所 東京都大田区下丸子 3 丁目 3 0 番 2 号

氏 名 キヤノン株式会社